

BUNDESREPUBLIK DEUTSCHLAND

09/831139

**PRIORITY
DOCUMENT**SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

REC'D	15 MAR 1999
WIPO	PCT

DE 99/68

EV

Bescheinigung

Die Siemens Aktiengesellschaft in München/Deutschland hat eine Patentanmeldung
unter der Bezeichnung

"Datenverarbeitungssystem oder Kommunikationsendgerät mit einer
Einrichtung zur Erkennung gesprochener Sprache und Verfahren
zur Erkennung bestimmter akustischer Objekte"

am 6. November 1998 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprüng-
lichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig das Symbol
G 10 L 15/22 der Internationalen Patentklassifikation erhalten.

München, den 2. Februar 1999

Deutsches Patent- und Markenamt**Der Präsident**

Im Auftrag

Aktenzeichen: 198 51 287.2

Beschreibung

Datenverarbeitungssystem oder Kommunikationsendgerät mit einer Einrichtung zur Erkennung gesprochener Sprache und Verfahren zur Erkennung bestimmter akustischer Objekte

Einrichtungen und Verfahren zur Erkennung natürlicher gesprochener Sprache sind dem Fachmann heute aus einer Vielzahl unterschiedlicher Anwendungen geläufig. Die praktische Anwendbarkeit und Leistungsfähigkeit derartiger Systeme hängt dabei stark von ihrer Komplexität und der Breite ihres Anwendungsbereiches ab. Generell gilt dabei, daß die Erkennungsrate eines solchen Systems mit steigender Zahl der zu erkennenden akustischen Objekte (Worte, Phoneme, Einzelbuchstaben, etc.) gewöhnlich stark abnimmt. Gleichzeitig steigt der Aufwand gemessen nach Kosten und Platzbedarf aber auch im Hinblick auf den Trainingsaufwand meist stark mit der Anwendungsbreite an.

Für viele Anwendungen kommen herkömmliche Spracherkennungssysteme deshalb noch immer nicht zum Einsatz, obwohl sie aus Sicht des Anwenders grundsätzlich dafür geeignet wären. Der Erfindung liegt daher die Aufgabe zugrunde, eine technische Lehre anzugeben, die den Einsatz der Spracherkennung auch für solche Anwendungen ermöglicht, bei denen ein größerer Aufwand sich aus wirtschaftlichen oder anderen Gründen verbietet. Diese Aufgabe wird durch ein Datenverarbeitungssystem oder Kommunikationsendgerät mit einer Einrichtung zur Erkennung gesprochener Sprache oder durch ein Verfahren zur Erkennung bestimmter akustischer Objekte nach einem der Patentansprüche gelöst.

Das erfindungsgemäße Erzeugnis, ein Datenverarbeitungssystem oder ein Kommunikationsendgerät, verfügt über eine Einrichtung zur Erkennung gesprochener Sprache, die speziell zur Erkennung bestimmter akustischer Objekte, nämlich von Einzelbuchstaben, Buchstabenkombinationen oder Steuerbefehlen eingerichtet oder speziell zur Erkennung solcher Objekte konfi-

zugte Ausführungsform z.B. die Ausgabe oder Anzeige von A02 nach der Ausgabe des falscherkannten Objekts A01, oder ähnliche Maßnahmen zur Unterstützung einer für den Benutzer möglichst komfortablen Korrektur des Erkennungsfehlers. Ein mögliche Wahl für ein solches besonderes akustisches Objekt oder einen solchen Steuerbefehl wäre z.B. das Wort "Falsch". Dem Fachmann fällt es nicht schwer, sich anhand der vorliegenden Darstellung weitere Anwendungsmöglichkeiten für diese Ausführungsform der vorliegenden Erfindung zu überlegen.

Weitere bevorzugte Ausführungsformen der vorliegenden Erfindung sind Gegenstand weiterer Unteransprüche.

Im folgenden wird die Erfindung anhand bevorzugter Ausführungsbeispiele und mit Hilfe von Figuren näher erläutert.

Figur 1 zeigt in schematischer Weise Aufbau und Funktionsweise einer bevorzugten Ausführungsform einer erfindungsgemäßen Vorrichtung.

Wie in Figur 1 dargestellt, umfaßt diese Ausführungsform eines erfindungsgemäßen Datenverarbeitungssystems (DPCD) oder Kommunikationsendgerätes (DPCD) eine Spracherkennungseinheit (SRU), die von einem Benutzer der Vorrichtung gesprochene akustische Objekte (AO) erkennt und die erkannten akustischen Objekte (RAO) einer Einrichtung zur akustischen Ausgabe oder optischen Anzeige (DU) zuführt. Gemäß der vorliegenden Erfindung ist die Spracherkennungseinrichtung speziell zur Erkennung bestimmter akustischer Objekte (AO), nämlich von Einzelbuchstaben, Buchstabenkombinationen oder Steuerbefehlen eingerichtet oder kann speziell zur Erkennung solcher Objekte konfiguriert werden.

Die Spracherkennungseinrichtung ordnet also jeweils einem vom Benutzer gesprochenen akustischen Objekt (AO) ein von dieser Einrichtung erkanntes akustisches Objekt (RAO) zu. Da die Erkennung natürlicher, gesprochener Sprache schon aus prinzipi-

Objekt in einem bestimmten Sinne nachfolgenden Objektes bewirkt. Die Folge der Objekte muß dabei nicht über die Größe von Erkennungswahrscheinlichkeiten oder Plausibilitätswerten festgelegt sein, sondern kann auch durch die Abfolge von Einträgen in einer Speichereinrichtung (MU) der Vorrichtung, oder durch alphabetische oder innerhalb eines definierten Kontextes semantisch definierte Abfolgen von Objekten gegeben sein. Zum Beispiel könnte die Abfolge der Objekte durch die Ordnung innerhalb einer Datenbank, eines Telefonbuches oder durch die Struktur einer in der Speichereinrichtung gespeicherten Datei, z.B. einer Kundendatei, eines Wörterbuches, oder ähnlicher Dateien gegeben sein.

Wenn in dieser Patentanmeldung von Einrichtungen die Rede ist, die für eine bestimmte Funktion oder Betriebsweise eingerichtet sind oder konfiguriert werden können, dann ist damit gemeint, daß die entsprechenden Funktionsmerkmale dieser Einrichtungen dauerhaft oder zeitlich begrenzt sein können. Ferner können diese Einrichtungen von allen Beteiligten zwischen Hersteller und Anwender durch Herstellprozesse, Einstellungen an der Hardware oder durch Verwendung oder Parametrierung einer Software oder gleichwirkende Mittel oder Maßnahmen für eine bestimmte Funktion oder Betriebsweise eingerichtet oder konfiguriert werden. Dem Fachmann erschließen sich anhand dieser Beschreibung ohne weiteres zahlreiche ähnliche oder gleichwertige Mittel oder Maßnahmen für diesen Zweck.

Eine Spracherkennungseinrichtung wird vorzugsweise durch eine geeignete Auswahl oder Parametrierung der Software eingerichtet oder konfiguriert, die die gewünschte Funktion im Spracherkennungsalgorithmus und oder die Ablaufsteuerung dieser Einrichtung realisiert. Ein Datenspeicher wird vorzugsweise durch eine geeignete Auswahl oder Parametrierung der Datenstruktur, z.B. der Datenbankstruktur, eingerichtet oder konfiguriert, die die Art der Speicherung der Daten auf diesem Speicher und die Art des Zugriffs auf diese Daten definiert.

erkennenden Objekte nach seiner Wahl (vorübergehend oder dauerhaft) auf eine bestimmte Untermenge wie z.B. Einzelbuchstaben, Buchstabenkombinationen oder Steuerbefehle einschränken kann.

5

Mit der vorliegenden Erfindung kann insbesondere die Anzahl der durch Sprachwahl abrufbaren Telefoneinträge in einem Mobiltelefon oder Schnurlostelefon oder in einem drahtgebundenen Telefon beliebig vergrößert werden. Bei herkömmlichen Systemen dieser Art wurde nur eine begrenzte Zahl von Einträgen für eine Sprachwahl zugelassen, erfahrungsgemäß maximal 20 oder 30 Einträge. Dies war durch den bereitzustellenden Speicherplatz für die wiederzuerkennenden Sprachsamples, d.h. durch die daraus entstehenden Kosten und den Platzbedarf bedingt. Wurde die Zahl der Einträge weiter erhöht, stieg erfahrungsgemäß der Aufwand für das Training der Spracherkennung erheblich, was zu einer geringeren Benutzerakzeptanz führte.

20 Gemäß einer bevorzugten Ausführungsform der vorliegenden Erfindung wird der Spracherkennungsalgorithmus vom Benutzer nur auf die Buchstaben des Alphabets und gegebenenfalls Kombinationen und einige wenige Steuerbefehle trainiert. Er wird auf diese Weise vom Benutzer zur Erkennung dieser akustischen Objekte eingerichtet oder entsprechend konfiguriert. Die Abfrage erfolgt durch akustische Eingabe von Anfangs- und (vorzugsweise bis zu zwei) Folgebuchstaben. Fehlerkennungen werden durch Plausibilitätsprüfungen, d.h. z.B. durch Vergleich der Objekte mit Einträgen in einer Speichereinrichtung, vermindert. Die eingegebenen Namen werden nur einmal
30 eingesprochen und in einem Coder mit geringer Bitrate (z.B. half-rate von GSM) gewandelt und unter dem entsprechenden Speicherplatz, gegebenenfalls komprimiert, gespeichert.

35 Alternativ kann auch ein Syntheseprogramm verwendet werden, das aus einem Namen Sprache synthetisiert, was eventuell weniger Speicherplatz erfordert. In jedem Fall muß die Spra-

Patentansprüche

1. Datenverarbeitungssystem (DPCD) oder Kommunikationsendge-
rät (DPCD) mit einer Einrichtung (SRU) zur Erkennung gespro-
chener Sprache mit folgenden Merkmalen:

a) die Einrichtung zur Spracherkennung ist speziell zur Er-
kennung bestimmter akustischer Objekte (AO), nämlich von Ein-
zelbuchstaben, Buchstabenkombinationen oder Steuerbefehlen
eingerrichtet oder kann speziell zur Erkennung solcher Objekte
konfiguriert werden;

b) es ist eine Einrichtung zur akustischen Ausgabe (DU) oder
optischen Anzeige (DU) erkannter akustischer Objekte (RAO)
vorgesehen.

15

2. Vorrichtung nach Anspruch 1, deren Einrichtung zur Spra-
cherkennung (SRU) so eingerichtet ist oder konfiguriert wer-
den kann, daß die Erkennung eines bestimmten ersten Steuerbe-
fehls im Anschluß an die Ausgabe oder Anzeige eines akusti-
schen Objektes die Ausgabe oder Anzeige eines weiteren aku-
stischen Objektes bewirkt.

20

3. Vorrichtung nach einem der vorhergehenden Ansprüche mit
einem Datenspeicher (MU), die so eingerichtet ist oder konfi-
guriert werden kann, daß die Erkennung eines akustischen Ob-
jekts oder einer Folge von Objekten das bzw. die einem Ein-
trag in dem Datenspeicher entspricht bzw. entsprechen, die
Anzeige oder Ausgabe dieses Eintrages (ME) oder eine mit die-
sem Eintrag verbundene Funktion (FU) der Vorrichtung bewirkt.

30

4. Vorrichtung nach Anspruch 3, bei der die Erkennungslei-
stung durch einen Vergleich möglicher Objekte oder Objektfol-
gen mit vorhandenen Einträgen im Datenspeicher (MU) verbes-
sert wird.

35

9. Verfahren nach einem der vorhergehenden Verfahrensansprüche, bei dem die Erkennungsleistung durch einen Vergleich möglicher Objekte oder Objektfolgen mit vorhandenen Einträgen im Datenspeicher verbessert wird.

5

10. Verfahren nach einem der vorhergehenden Verfahrensansprüche, dessen Spracherkennungsalgorithmus mit Hilfe bestimmter Steuerbefehle in spezielle Betriebszustände zur Erkennung von Einzelbuchstaben, Buchstabenkombinationen oder Steuerbefehlen gebracht werden kann.

10

1/1

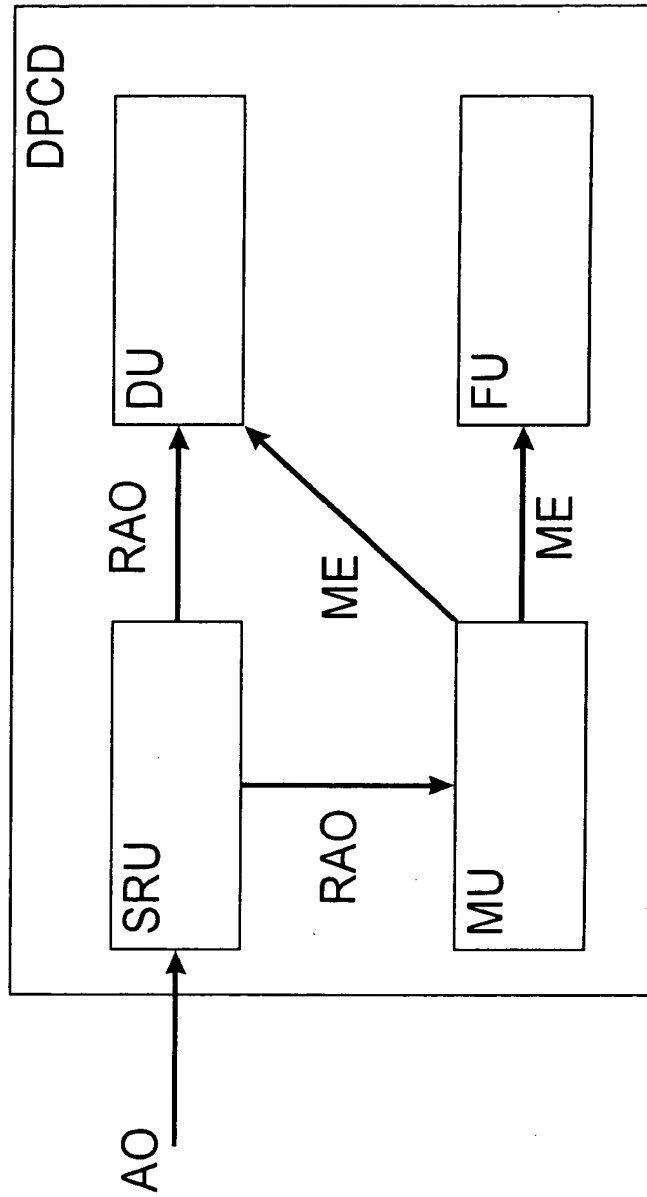


Fig. 1